

Age Optimal Information Gathering and Dissemination on Graphs

Vishrant Tripathi, Rajat Talak, and Eytan Modiano
Laboratory for Information & Decision Systems, MIT

Abstract—We consider the problem of timely exchange of updates between a central station and a set of ground terminals V , via a mobile agent that traverses across the ground terminals along a mobility graph $G = (V, E)$. We design the trajectory of the mobile agent to minimize peak and average age of information (AoI), two newly proposed metrics for measuring timeliness of information. We consider randomized trajectories, in which the mobile agent travels from terminal i to terminal j with probability $P_{i,j}$. For the information gathering problem, we show that a randomized trajectory is peak age optimal and factor- $8\mathcal{H}$ average age optimal, where \mathcal{H} is the mixing time of the randomized trajectory on the mobility graph G . We also show that the average age minimization problem is NP-hard. For the information dissemination problem, we prove that the same randomized trajectory is factor- $O(\mathcal{H})$ peak and average age optimal. Moreover, we propose an age-based trajectory, which utilizes information about current age at terminals, and show that it is factor-2 average age optimal in a symmetric setting.

I. INTRODUCTION

Many emerging applications depend on the collection and delivery of status updates between a set of ground terminals and a central terminal using mobile agents. Examples include: measuring traffic at road intersections [1], temperature, and pollution in cities [2], ocean monitoring using underwater autonomous vehicles [3], and surveillance using UAVs [4]. All of these applications depend upon regular status updates, to be communicated in a timely manner, so as to keep the central terminal and the ground terminals updated with fresh information.

Age of Information (AoI) is a newly proposed metric that captures timeliness of the received information [5]–[7]. Unlike packet delay, AoI measures the lag in obtaining information at the destination node, and is therefore suited for applications involving gathering or dissemination of time sensitive updates. Age of information, at a destination, is defined as the time that elapsed since the last received information update was generated at the source. AoI, upon reception of a new update packet, drops to the time elapsed since generation of the packet, and grows linearly otherwise.

We consider the problem of AoI minimization in gathering and dissemination of information updates, between a set of ground terminals and a central terminal. The information updates can be as small as a single packet containing temperature information or a high fidelity image or a video file. The ground terminals are equipped with low power transmitters, and a mobile agent is used to gather and disseminate information.

This work was supported by NSF Grants AST-1547331, CNS-1713725, and CNS-1701964, and by Army Research Office (ARO) grant number W911NF-17-1-0508.

The age or freshness of information gathered and disseminated depends on the trajectory of the mobile agent, whose mobility is constrained to a *mobility graph* $G = (V, E)$. The mobile agent can move from ground terminal i to ground terminal j only if $(i, j) \in E$. This model can be used to capture the fact that the agent may not be able to move between any arbitrary locations due to topological limitations.

The problem of persistent monitoring in dynamic environments has been considered in [8] and [9] using tools from optimal control. These works focus on minimizing uncertainty when source locations are time varying, rather than timely monitoring over a fixed set of locations. Minimizing delay in a similar setting with packets arriving randomly in space and time has been considered in [10].

Closer to our work is [11] in which some approximation trajectories to minimize maximum latency on metric graphs were proposed. In [12], the authors consider trajectory planning for a mobile agent to minimize AoI. They obtain the best permutation of nodes for the mobile agent to visit in sequence, given Euclidian distances between the nodes. In our work, mobility is constrained by a general graph G , and we seek the optimal trajectory over the space of all trajectories allowed on this graph G , not just permutations of nodes. To the best of our knowledge, this is the first work to consider the AoI minimization on general mobility graphs G , and provide polynomial time approximation algorithms.

In the information gathering problem, we consider the design of trajectories for the mobile agent to minimize peak and average age, two popular metrics of AoI. We first consider the space of randomized trajectories, in which the mobile agent traverses edges according to a random walk on the mobility graph G . We show that a randomized trajectory is in fact peak age optimal, and that it can be obtained in polynomial time using the Metropolis-Hastings algorithm. We then prove that solving for the average age optimal trajectory is NP-hard, in a symmetric setting, and propose a heuristic randomized trajectory that is simultaneously peak age optimal and factor- $8\mathcal{H}$ average age optimal, where \mathcal{H} is the mixing time of the randomized trajectory on G . The factor \mathcal{H} can scale with the graph size, especially if the graph is not well connected. Thus, we propose an age-based trajectory, in which the mobile agent uses the current AoI to determine its motion, and show that it is factor-2 optimal in a symmetric setting.

In the information dissemination problem, the central terminal sends updates for each ground terminal via the mobile agent. The mobile agent queues these update packets in a

first-come-first-serve (FCFS) queue, and delivers them to the respective ground terminal when the mobile agent reaches it. The FCFS queue assumption is motivated by uncontrollable MAC layer queues, where the generated updates get queued for transmission [7], [13]. We, now, not only have to design the trajectory of the mobile agent, but also determine update generation rate for each ground terminal. We show that the peak age optimal randomized trajectory of the information gathering problem, along with a simple update generation rate, is at most a factor- $O(\mathcal{H})$ optimal, in both peak and average age. Also derived is an explicit formula for peak age of the discrete time Ber/G/1 queue with vacations.

We describe the system model in Section II. The information gathering and dissemination problems are studied in Section III and Section IV, respectively. We present simulation results in Section V, and conclude in Section VI. We have omitted some of the proofs due to space constraints. See the technical report for details [14].

II. SYSTEM MODEL

We consider a central terminal that needs to communicate with a set of ground terminals V . The ground terminals are equipped with low power, low range radio communication devices, and cannot directly communicate with the central terminal, or with each other. An autonomous mobile agent m , is used as a relay between the central terminal and the ground terminals, while moving across the geographical region where the ground terminals are spread.

The mobility of the agent is constrained by a *mobility graph* $G = (V, E)$, where m can travel from ground terminal i to ground terminal j only if $(i, j) \in E$. The graph G , thus, constraints the set of allowable moves. We consider a time-slotted system, with slot duration normalized to unity. In the duration of a time-slot, the mobile agent stays at a ground terminal to gather or disseminate information, and moves to any of its neighbours in G for the next time-slot. The mobility graph can be constructed from the limitations of a slot duration, distances between ground terminals, and speed of the mobile agent.

We consider two problems: *information gathering* and *information dissemination*. In the information gathering problem, every time the mobile agent reaches a ground terminal $i \in V$, the ground terminal sends a fresh update to the mobile agent, which is immediately relayed to the central terminal. The age $A_i(t)$, at the central terminal, for the ground terminal i drops to 1. When the mobile agent is not at the ground terminal i , the age $A_i(t)$ increases linearly; see Figure 1. The evolution of $A_i(t)$ in the information gathering problem is given by:

$$A_i(t+1) = \begin{cases} A_i(t) + 1, & \text{if } m(t) \neq i \\ 1, & \text{if } m(t) = i \end{cases} \quad (1)$$

where $m(t)$ denotes the location of the mobile agent at time t . Note that the age evolution depends on the trajectory that the mobile agent follows on the mobility graph G .

In the information dissemination problem, the central terminal generates updates for each ground terminal. The generated

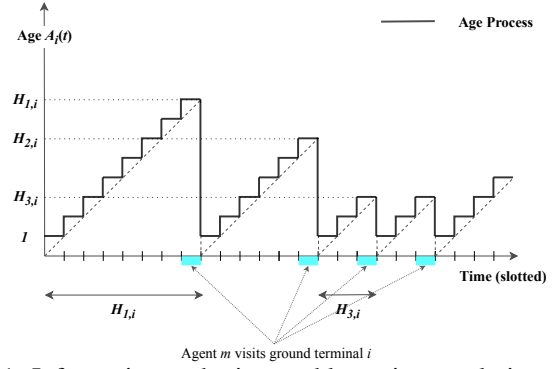


Fig. 1: Information gathering problem: time evolution of age $A_i(t)$; $H_{k,i}$ is the k^{th} return time to terminal i .

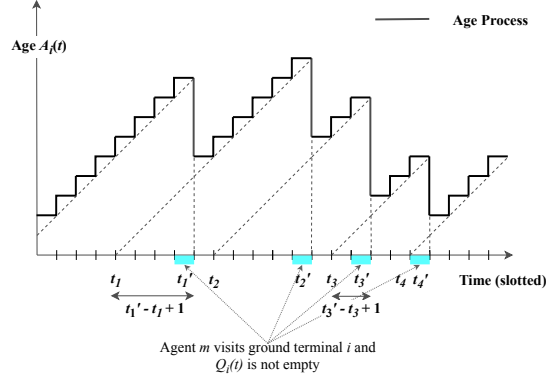


Fig. 2: Information dissemination problem: time evolution of age $A_i(t)$; t_k, t'_k are the generation and reception times of the k^{th} status update for terminal i .

updates are then transmitted to the mobile agent. The mobile agent queues updates received from the central terminal in a set of V FCFS queues, one for each ground terminal. The mobile agent delivers the head-of-line update in queue i , to ground terminal i , when it reaches i . The central terminal has no control over the FCFS queues on the mobile agent, however, can control the update generation rate λ_i , for each ground terminal i .

The age $A_i(t)$, at the ground terminal i , increases by 1 every time the mobile agent is not at i , or when it is at i but has no updates to deliver. Otherwise, a successful delivery of the head-of-line update occurs in time slot t , and the age $A_i(t)$ drops to the age of the head-of-line update in queue i . See Figure 2. This evolution of age $A_i(t)$ can be written as:

$$A_i(t+1) = \begin{cases} A_i(t) + 1, & \text{if } m(t) \neq i \\ A_i(t) + 1, & \text{if } m(t) = i \text{ and } Q_i(t) = \emptyset \\ t - G_i(t) + 1, & \text{if } m(t) = i \text{ and } Q_i(t) \neq \emptyset \end{cases} \quad (2)$$

where $G_i(t)$ is the time of generation of the head of line packet in queue i , at time t , and $Q_i(t)$ denotes the set of packets in the mobile agent's queue i at time t .

A. Age Metrics

AoI is an evolving function of time. We consider two time average metrics of AoI. Average age, for ground terminal i ,

is defined as the time averaged area under the age curve:

$$A_i^{\text{ave}} \triangleq \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T A_i(t). \quad (3)$$

In Figures 1 and 2, we see that the age $A_i(t)$ peaks before a fresh update is delivered. In the information gathering case, a fresh update is delivered every time the mobile agent visits i , i.e. $m(t) = i$. Whereas, in the information dissemination case, a fresh update is delivered whenever $m(t) = i$ and the queue $Q_i(t) \neq \emptyset$. The peak age A_i^p , for ground terminal i , defined as an average of all the peaks in the age evolution curve $A_i(t)$, can be written as

$$A_i^p \triangleq \limsup_{T \rightarrow \infty} \frac{\sum_{t=1}^T A_i(t) \mathbb{1}_{\{m(t)=i\}}}{\sum_{t=1}^T \mathbb{1}_{\{m(t)=i\}}}, \quad (4)$$

in the information gathering case and

$$A_i^p \triangleq \limsup_{T \rightarrow \infty} \frac{\sum_{t=1}^T A_i(t) \mathbb{1}_{\{m(t)=i, Q_i(t) \neq \emptyset\}}}{\sum_{t=1}^T \mathbb{1}_{\{m(t)=i, Q_i(t) \neq \emptyset\}}}, \quad (5)$$

in the information dissemination case.

We define the network peak and average age to be

$$A^p = \sum_{i \in V} w_i A_i^p \quad \text{and} \quad A^{\text{ave}} = \sum_{i \in V} w_i A_i^{\text{ave}}, \quad (6)$$

where $w_i > 0$ are weights representing the relative importance of a ground terminal i . Our goal is to minimize network peak and average age.

B. Trajectory Space

We use \mathbb{T} to denote a reasonably large space of trajectories:

$$\mathbb{T} = \{ \text{Trajectory } \mathcal{T} \mid f_i(\mathcal{T}) \text{ exists and is positive } \forall i \in V \},$$

where $f_i(\mathcal{T})$ denotes the fraction of time-slots, the trajectory \mathcal{T} , is at ground terminal i :

$$f_i(\mathcal{T}) = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{1}_{\{m(t)=i\}}. \quad (7)$$

For a trajectory $\mathcal{T} \in \mathbb{T}$, the limit (7) exists and is positive for all $i \in V$. This requirement is to ensure that the peak and average age are both finite and well defined.

Peak and average age depend on the trajectory $\mathcal{T} \in \mathbb{T}$. We use $A^p(\mathcal{T})$ and $A^{\text{ave}}(\mathcal{T})$ to denote network peak and average age, respectively, for $\mathcal{T} \in \mathbb{T}$.

III. INFORMATION GATHERING

In this section, we consider the problem of information gathering. We define optimal peak and average age to be

$$A_G^p = \min_{\mathcal{T} \in \mathbb{T}} A^p(\mathcal{T}), \quad \text{and} \quad A_G^{\text{ave}} = \min_{\mathcal{T} \in \mathbb{T}} A^{\text{ave}}(\mathcal{T}), \quad (8)$$

where \mathbb{T} denotes the space of all trajectories for the mobile agent.

We first consider randomized trajectories, where the mobile agent moves according to a random walk on the mobility graph. We shall show that for peak age optimality, such randomized trajectories suffices. We then show that the average age optimization is NP-hard, and propose a heuristic randomized trajectory. In Section III-D, we propose an age-based trajectory for better average age performance.

A. Randomized Trajectories

We start by defining the class of randomized trajectories:

Definition A trajectory $m(t)$, on mobility graph G , is said to be a *randomized trajectory* if $m(t)$ is an irreducible Markov chain defined by a transition probability matrix \mathbf{P} :

$$\mathbb{P}[m(t+1) = j | m(t) = i] = P_{i,j}, \quad (9)$$

for all t and $i, j \in V$, where $P_{i,j} = 0$ for $(i, j) \notin E$.

For convenience, we shall refer to $m(t)$, defined above, as the randomized trajectory \mathbf{P} , where \mathbf{P} to denote the matrix with entries $P_{i,j}$. Note that $P_{i,j}$ is the probability that the mobile agent, when at ground terminal i , moves to ground terminal j for the next time slot. The constraint: $P_{i,j} = 0$ for $(i, j) \notin E$, ensures that the randomized trajectory adheres to the mobility constraints defined by G .

We assume in the definition of a randomized trajectory \mathbf{P} , that $m(t)$ is an irreducible Markov chain over the state space V . This is desired, since the mobile agent has to visit each ground terminal for a positive fraction of time, or otherwise the resulting peak and average age would be unbounded.

For any randomized trajectory \mathbf{P} , we obtain explicit expressions for network peak and average age. We use the notation $A^p(\mathbf{P})$ and $A^{\text{ave}}(\mathbf{P})$ to show explicit dependence of peak and average age on the randomized trajectory \mathbf{P} .

Theorem 1: The network peak and average age for a randomized trajectory \mathbf{P} is given by

$$A^p(\mathbf{P}) = \sum_{i \in V} \frac{w_i}{\pi_i}, \quad \text{and} \quad A^{\text{ave}}(\mathbf{P}) = \sum_{i \in V} \frac{w_i z_{ii}}{\pi_i}, \quad (10)$$

where π is the unique stationary distribution obtained by solving $\pi \mathbf{P} = \pi$ and z_{ii} are diagonal elements of the matrix $Z \triangleq (I - \mathbf{P} + \Pi)^{-1}$, where Π is an $n \times n$ matrix with entries $\Pi_{i,j} \triangleq \pi_j$, $\forall i, j \in V$.

Proof: From Figure 1, observe that the l th peak for $A_i(t)$ equals the l th return times $H_{l,i}$, and therefore, the peak age $A_i^p = \mathbb{E}[H_{1,i}]$. Since the age curve $A_i(t)$ increases in steps from 1 to $H_{l,i}$, for every l th return time, the average age A_i^{ave} can be computed to be $(\mathbb{E}[H_{1,i}^2] + \mathbb{E}[H_{1,i}]) / (2\mathbb{E}[H_{1,i}])$.

For a randomized trajectory \mathbf{P} , the first and second moment of return times are given by $\mathbb{E}[H_{1,i}] = \frac{1}{\pi_i}$ and $\mathbb{E}[H_{1,i}^2] = \frac{1}{\pi_i} + \frac{2z_{ii}}{\pi_i^2}$, respectively; see [15]. Substituting this we can obtain the

explicit expressions for peak and average age. A detailed proof is in our technical report [14]. ■

B. Peak Age Minimization

We first formulate the peak age minimization problem over the space of randomized trajectories. We shall see that a peak age optimal randomized trajectory suffices for optimality over the space of all trajectories.

Using Theorem 1, we can write the peak age minimization problem over the space of randomized trajectories as:

$$\begin{aligned} & \text{Minimize}_{\mathbf{P}, \pi} \quad \sum_{i \in V} \frac{w_i}{\pi_i}, \\ & \text{subject to} \quad P_{i,j} \geq 0, \forall (i,j), \text{ and } \mathbf{P}\mathbf{1} = \mathbf{1}, \\ & \quad \pi\mathbf{P} = \pi, \mathbf{1}^T \pi = 1, \text{ and } \pi_i \geq 0 \forall i \\ & \quad P_{i,j} = 0, \forall (i,j) \notin E, \\ & \quad \mathbf{P} \text{ is irreducible.} \end{aligned} \quad (11)$$

Here, \mathbf{P} is a randomized trajectory, while π is the unique stationary distribution associated with it. This problem is difficult to solve because the irreducibility constraint cannot be expressed in a simple, solvable manner. Further, relaxing the irreducibility constraint can yield a trivial solution like $\mathbf{P} = I$, which are neither irreducible nor close to optimal.

However, the problem (11) can be transformed to finding an irreducible \mathbf{P} , with a given stationary distribution. This is a simpler problem and can be solved using the Metropolis-Hastings algorithm.

Lemma 1: Let $\pi_i^* \triangleq \frac{\sqrt{w_i}}{\sum_{j \in V} \sqrt{w_j}}$, for all $i \in V$, to be a distribution on V , and a randomized trajectory \mathbf{P} satisfy $\pi^* \mathbf{P} = \pi^*$. Then, (π^*, \mathbf{P}) solves (11).

Proof: If we were to minimize the objective in (11) over the space of all distribution π on V , then the optimal distribution would be $\pi_i^* = \sqrt{w_i} / \sum_{j \in V} \sqrt{w_j}$, for all $i \in V$.

Thus, if a randomized trajectory \mathbf{P} exists which induces this distribution on V , then clearly (π^*, \mathbf{P}) would solve (11). ■

Lemma 1 implies that a randomized trajectory \mathbf{P} , that satisfies $\pi^* \mathbf{P} = \pi^*$, is a peak age optimal, over the space of all randomized trajectories. We now construct one such randomized trajectory: for π^* given in Lemma 1, define a Metropolis-Hastings randomized trajectory \mathbf{P}^{mh} :

$$P_{i,j}^{\text{mh}} = \begin{cases} \frac{1}{d_i} \min(1, \frac{\pi_j^* d_j}{\pi_i^* d_i}), & \text{if } i \neq j \text{ and } (i,j) \in E \\ 1 - \sum_{j:j \neq i} P_{i,j}^{\text{mh}}, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}, \quad (12)$$

where d_i equals the out degree of terminal i in the mobility graph G . It is known that such a randomized trajectory \mathbf{P}^{mh} satisfies $\pi^* \mathbf{P} = \pi^*$ [15]. This gives the following result.

Theorem 2: The Metropolis-Hastings randomized trajectory \mathbf{P}^{mh} solves (11), i.e. it is peak age optimal over the space of all randomized trajectories.

Thus far, we considered randomized trajectories. We now show that, for peak age optimality, this randomization suffices.

Theorem 3: The Metropolis-Hastings randomized trajectory \mathbf{P}^{mh} is peak age optimal over the space of all trajectories \mathbb{T} , namely $A^{\text{p*}}(\mathbf{P}^{\text{mh}}) = A_G^{\text{p*}}$.

Proof: We establish a more general result. Namely, any randomized trajectory which satisfies $\pi^* \mathbf{P} = \pi^*$ is peak age optimal over the space of all trajectories, where π^* is given in Lemma 1. See Appendix A. ■

This result establishes that the peak age minimization can be solved in polynomial time, as we are able to compute an optimal trajectory \mathbf{P}^{mh} in polynomial time; in $O(|V|^2)$ time.

C. Average Age Minimization

We now consider the average age minimization problem. We prove that in the symmetric setting, namely $w_i = 1 \forall i \in V$,¹ the average age minimization problem is NP-hard. The proof essentially establishes an equivalence between the average age minimization problem and the Hamiltonian cycle problem. See [14] for details.

Since solving the average age minimization problem is hard, we derive a lower bound on average age. Intuitively, if the mobility graph is better connected then it should yield a lower age. This is because a better connected mobility graph imposes fewer restrictions on mobility. The following result obtains a lower bound on network average age by comparing it with the network average age of a complete graph.

Theorem 4: For any trajectory $\mathcal{T} \in \mathbb{T}$, the network average age is lower bounded by

$$A^{\text{ave}}(\mathcal{T}) \geq \frac{1}{2} \sum_{i \in V} \left(\frac{w_i}{\pi_i^*} + w_i \right), \quad (13)$$

where $\pi_i^* = \frac{\sqrt{w_i}}{\sum_{j \in V} \sqrt{w_j}}$ for all $i \in V$.

Proof: We use the Cauchy-Schwarz inequality along with the peak age result to lower bound the network average age on a complete graph. The proof is in our technical report [14]. ■

Note that the term $\sum_{i \in V} \frac{w_i}{\pi_i^*}$ is nothing but the optimal peak age $A_G^{\text{p*}}$; see Theorem 3. Also, the lower bound in Theorem 4 is independent of the trajectory \mathcal{T} . Thus, we have

$$A_G^{\text{ave*}} = \min_{\mathcal{T} \in \mathbb{T}} A^{\text{ave}}(\mathcal{T}) \geq A_{\text{LB}}^{\text{ave}} = \frac{1}{2} A_G^{\text{p*}} + \frac{1}{2} \sum_{i \in V} w_i, \quad (14)$$

¹The weights w_i only measure relative significance of ground terminals. Thus, setting $w_i = 1 \forall i \in V$ is equivalent to setting $w_i = w_j \forall i, j \in V$.

where \mathbb{T} is the space of all trajectories. It must be noted that a similar result was derived in the case of link scheduling for age minimization in [7].

1) *A Heuristic Randomized Trajectory*: Motivated by the peak age optimality results of the previous section, we restrict ourselves to the space of randomized trajectories, and propose a heuristic, called the *fastest-mixing randomized trajectory*, and prove an average age performance bound for it.

Using the results in Theorem 1, the average age minimization problem over the space of randomized trajectories can be written as

$$\begin{aligned} & \underset{\mathbf{P}, \pi, \mathbf{Z}}{\text{Minimize}} && \sum_{i \in V} \frac{w_i z_{ii}}{\pi_i}, \\ & \text{subject to} && P_{i,j} \geq 0, \forall (i, j), \text{ and } \mathbf{P}\mathbf{1} = \mathbf{1}, \\ & && \pi\mathbf{P} = \pi, \mathbf{1}^T \pi = 1, \text{ and } \pi_i \geq 0 \forall i \\ & && P_{i,j} = 0, \forall (i, j) \notin E, \\ & && \mathbf{P} \text{ is irreducible,} \\ & && \Pi_{i,j} = \pi_j \forall (i, j), \\ & && \mathbf{Z} = (\mathbf{I} - \mathbf{P} + \Pi)^{-1}. \end{aligned} \quad (15)$$

Here, \mathbf{P} is the randomized trajectory, and π the unique stationary distribution associated with it. Solving (15) can be computationally complex due to the irreducibility and the non-linearity constraint in $\mathbf{Z} = (\mathbf{I} - \mathbf{P} + \Pi)^{-1}$.

We next upper bound the network average age, for any randomized trajectory \mathbf{P} of the mobile agent. We first define mixing time for a randomized trajectory.

To do this, we first discuss the notion of stopping rules and stopping times in a Markov chain. A stopping rule is a rule that observes the walk on a Markov chain and, at each step, decides whether or not to stop the walk based on the steps so far. Stopping rules can make probabilistic decisions and therefore the time at which the walk stops, called the stopping time, is a random variable.

Mixing Time [16] The hitting time from state distribution σ_1 to σ_2 on a Markov chain is the minimum expected stopping time over all stopping rules that, beginning at σ_1 , stop in the exact distribution of σ_2 . This is denoted by $\mathcal{H}(\sigma_1, \sigma_2)$. The mixing time \mathcal{H} of a Markov chain \mathbf{P} is then defined as

$$\mathcal{H} \triangleq \sup_{\sigma \in \Delta(V)} \mathcal{H}(\sigma, \pi), \quad (16)$$

where $\Delta(V)$ is the collection of all distributions on V and π is the stationary distribution of \mathbf{P} .

Lemma 2: The network average age for a randomized trajectory \mathbf{P} is upper bounded by

$$A^{\text{ave}}(\mathbf{P}) = \sum_{i \in V} \frac{w_i z_{ii}}{\pi_i} \leq 4\mathcal{H}A^{\text{P}}(\mathbf{P}) + \sum_{i \in V} w_i, \quad (17)$$

where \mathcal{H} denotes the mixing time of the randomized trajectory \mathbf{P} .

Proof: Define $\mathcal{Z} \triangleq \max_i \sum_j |z_{ij} - \pi_j|$, known as the discrepancy for a randomized trajectory \mathbf{P} . This definition implies that $z_{ii} \leq \mathcal{Z} + \pi_i, \forall i \in V$. Thus, we get the following upper bound:

$$\sum_{i \in V} \frac{w_i z_{ii}}{\pi_i} \leq \sum_{i \in V} \left(\frac{w_i \mathcal{Z}}{\pi_i} + w_i \right).$$

However, from [17] we know that $\mathcal{Z} \leq 4\mathcal{H}$, where \mathcal{H} is the mixing time of the randomized trajectory \mathbf{P} . Thus, we have the required result

$$\sum_{i \in V} \frac{w_i z_{ii}}{\pi_i} \leq \sum_{i \in V} \left(\frac{4w_i \mathcal{H}}{\pi_i} + w_i \right) = 4\mathcal{H}A^{\text{P}}(\mathbf{P}) + \sum_{i \in V} w_i,$$

where the last equality follows from Theorem 1. ■

We use this relation and suggest the following heuristic for minimizing age: *Find the fastest mixing randomized trajectory \mathbf{P} on the mobility graph G that minimizes peak age.*

From the proof of Theorem 3, we know that for a randomized trajectory \mathbf{P} to be peak age optimal all we need is $\pi_i \propto \sqrt{w_i}$, where π is the stationary distribution of \mathbf{P} . It, therefore, suffices to find \mathbf{P} that satisfies $\pi_i \propto \sqrt{w_i}$, and simultaneously minimizes the mixing time \mathcal{H} . We call this the *fastest-mixing randomized trajectory*, and use \mathbf{P}^* to denote it. The following result provides a way to obtain \mathbf{P}^* by solving a convex program.

Theorem 5: The fastest mixing randomized trajectory can be found by solving the following convex optimization problem:

$$\begin{aligned} & \underset{\mathbf{P}}{\text{Minimize}} && \mu(\mathbf{P}) = \|\mathbf{P} - \Pi^*\|_2, \\ & \text{subject to} && P_{i,j} \geq 0, \forall (i, j), \\ & && \mathbf{P}\mathbf{1} = \mathbf{1}, \\ & && \pi^* \mathbf{P} = \pi^*, \quad \Pi_{i,j}^* = \pi_i^* \forall i, j \in V, \\ & && P_{i,j} = 0, \forall (i, j) \notin E. \end{aligned} \quad (18)$$

Here $\|A\|_2$ denotes the spectral norm of matrix A and $\pi_i^* = \frac{\sqrt{w_i}}{\sum_{j \in V} \sqrt{w_j}}, \forall i \in V$.

Proof: See Appendix B. ■

This convex program (18) finds a randomized trajectory \mathbf{P} on G that is closest to the stationary randomized walk Π^* , in the spectral norm sense. Also, \mathbf{P}^* is peak age optimal on graph G , since it satisfies $\pi_i^* \propto \sqrt{w_i}$. Note that, the problem (18) can be solved in polynomial time by converting it to a semi-definite program [18].

We now bound the average age performance of the fastest-mixing randomized trajectory.

Theorem 6: The network average age of the fastest-mixing randomized trajectory is at most $8\mathcal{H}$ -factor away

from the optimal average age:

$$\frac{A^{\text{ave}}(\mathbf{P}^*)}{A_G^{\text{ave}*}} \leq 8\mathcal{H}, \quad (19)$$

where \mathcal{H} is the mixing time of P^* .

Proof See Appendix C.

To see the usefulness of the fastest-mixing randomized trajectory, and Theorem 6, consider a random geometric graph $\mathcal{G}(n, r)$. The graph consists of n nodes spread over a unit square with a link between every two nodes that are within a distance r . If v is the physical speed of the mobile agent, then r must equal $v\tau$, where τ is the slot duration. We know that mixing time of $\mathcal{G}(n, r)$ is $O\left(\frac{\log n}{r^2}\right)$, and therefore, the fastest-mixing randomized trajectory would be at most $O\left(\frac{\log n}{v_{\text{max}}^2 \tau^2}\right)$ factor optimal. For highly connected graphs, such as Dirac graphs in which the degree of each node is at least $|V|/2$, we have constant factor of optimality; since the mixing times are $O(1)$ [19].

D. Age-based Trajectories

In the last two sub-sections, we proposed two randomized trajectories, namely \mathbf{P}^{mh} and \mathbf{P}^* . Both were peak age optimal, while the latter was also factor- \mathcal{H} average age optimal. We also noted that solving the average age problem is generally hard. We now propose an age-based trajectory which can be constant factor age optimal.

Age-based trajectory In every time slot, agent m moves to the location that has the highest weighted function of $A_i(t)$. Specifically, if $m(t) = i$ then

$$m(t+1) = \arg \max_{j:(i,j) \in E} w_j g(A_j(t)), \quad (20)$$

for all $i, j \in V$ and time t , where $g(\cdot)$ is an increasing function.

In the symmetric setting, where $w_i = 1 \forall i \in V$, we prove that the age-based trajectory is factor-2 optimal.

Theorem 7: In the symmetric setting $w_i = 1 \forall i \in V$, the network average age A^{ave} for the age-based trajectory is bounded by

$$\frac{A^{\text{ave}}}{A_G^{\text{ave}*}} \leq \frac{2|V|+1}{|V|+1} \leq 2, \quad (21)$$

for any increasing function $g(\cdot)$.

Proof: In the symmetric setting, we observe that the age-based trajectory does a repeated depth-first search (DFS) of graph G . In a DFS traversal, we note that every location gets visited at least once in every $2|V|$ time-slots. This is because,

every vertex is visited at least once and visits after the first visit to all nodes is at most $|V|$. This is because every repeated visit to a vertex means that at least one new vertex was visited.

Thus, the average age can be upper bounded as: $A_i^{\text{ave}} \leq (2|V|+1)/2$ and $A^{\text{ave}} \leq |V|(2|V|+1)/2$. From Lemma 4, we obtain the average age lower-bound to be $A_{\text{LB}}^{\text{ave}} = |V|(|V|+1)/2$. Combining this, we get $A^{\text{ave}}/A_{\text{LB}}^{\text{ave}} \leq (2|V|+1)/(|V|+1)$. The proof follows because $A_{\text{LB}}^{\text{ave}} \leq A_G^{\text{ave}*}$, since $A_{\text{LB}}^{\text{ave}}$ is a lower-bound. ■

IV. INFORMATION DISSEMINATION

We now consider the information dissemination problem. The central terminal generates updates for every ground terminal i , at rate λ_i , according to a Bernoulli process. The generated updates for the ground terminal i are sent to the mobile agent, which get queued in the i th FCFS queue. The mobile agent follows a trajectory \mathcal{T} , and delivers the head-of-line update in queue i to terminal i , when it reaches it.

Our objective is to minimize the network peak age and average age over the space of update generation rates λ and all trajectories \mathcal{T} :

$$A_D^{\text{p}*} = \min_{\mathcal{T} \in \mathbb{T}, \lambda} \sum_{i \in V} w_i A_i^{\text{p}}, \quad \text{and} \quad A_D^{\text{ave}*} = \min_{\mathcal{T} \in \mathbb{T}, \lambda} \sum_{i \in V} w_i A_i^{\text{ave}}, \quad (22)$$

where A_i^{p} denotes peak age and A_i^{ave} denotes the average age of terminal i . Their evolution is given by (2). For convenience, we have omitted their explicit dependence on $\mathcal{T} \in \mathbb{T}$ and λ .

Motivated by the results for the information gathering problem, we consider randomized trajectories. Note that an arriving update in queue i has service time equal to the inter-visit times to ground terminal i , provided the update arrived when the queue i was not-empty; $Q_i(t) \neq \emptyset$. However, when an update arrives to an empty queue i , the time to delivery is not the inter-visit time, and depends on the location of the mobile agent at the time of arrival.

Since the analysis of age for such a queueing system may be difficult, we provide an upper bound, by comparing the i th queue with a discrete time Ber/G/1 queue with vacations: whenever the i th queue is empty pretend that it goes on a vacation, with vacation times having the same distribution as inter-visit time; otherwise the service times for the queue are just inter-visit times. Clearly, the age process of such a FCFS queue is an upper bound for the age process $A_i(t)$. Thus, we upper bound the peak age A_i^{p} and average age A_i^{ave} , by the peak and average age of this Ber/G/1 queue with vacations. We first analyze peak and average age of a Ber/G/1 queue with vacations.

A. Age for Ber/G/1 Queue with Vacations

Consider a discrete time FCFS Ber/G/1 queue with vacations, where an arrival occurs with probability λ , the service times S are generally distributed with mean $\mathbb{E}[S] = 1/\mu$, and the vacation times V are also generally distributed. We obtain an expression for the peak age of a discrete time Ber/G/1 queue with vacations, and a bound on average age.

Lemma 3: The peak age for a discrete time FCFS Ber/G/1 queue with vacations is given by

$$A^p = \frac{1}{\lambda} + \frac{1}{\mu} + \frac{\lambda \mathbb{E}[S^2] - \rho}{2(1-\rho)} + \frac{\mathbb{E}[V^2]}{2\mathbb{E}[V]} - \frac{1}{2}, \quad (23)$$

where $\rho = \frac{\lambda}{\mu}$, while the average age is upper-bounded by peak age, namely $A^{\text{ave}} \leq A^p$.

Proof: Proof is in the technical report [14]. ■

B. Age Minimization Problem

Using Lemma 3, we now obtain an upper-bound on both, network peak and average age, for a given randomized trajectory \mathbf{P} and update generation rates $\boldsymbol{\lambda}$.

Lemma 4: For a randomized trajectory \mathbf{P} and packet generation rates $\boldsymbol{\lambda}$, the peak and average age for a ground terminal i is upper-bounded by

$$A_i^{\text{UB}} = \frac{1}{\pi_i} \left[1 + z_{ii} + \frac{1}{\rho_i} + \frac{z_{ii}\rho_i}{1-\rho_i} \right] - \frac{\rho_i}{1-\rho_i} - 1, \quad (24)$$

for all $i \in V$, where π is the unique stationary distribution of \mathbf{P} , $Z = (I - \mathbf{P} + \Pi)^{-1}$, Π is a matrix with all rows equal to the stationary distribution vector π , and $\rho_i \triangleq \frac{\lambda_i}{\pi_i}$.

Proof: This is obtained by substituting S and V in Lemma 3 by the inter-visit times $H_{1,i}$, for queue i . We relegate the details to our technical report [14]. ■

We propose a policy, i.e. a randomized trajectory \mathbf{P} and update generation rate $\boldsymbol{\lambda}$, that minimizes the age upper-bound $A^{\text{UB}} = \sum_{i \in V} w_i A_i^{\text{UB}}$:

Definition Separation Principle Policy

- 1) Mobile agent follows the randomized trajectory \mathbf{P}^* obtained by solving (18).
- 2) Generate updates for the ground terminal i at rate

$$\lambda_i^* = \frac{\pi_i^*}{1 + \sqrt{z_{ii}^* - \pi_i^*}}, \quad (25)$$

where $\pi_i^* = \frac{\sqrt{w_i}}{\sum_{j \in V} w_j}$ and z_{ii} are diagonal elements of the matrix $Z = (I - \mathbf{P}^* + \Pi^*)^{-1}$.

We call it the separation principle policy for two reasons. Firstly, \mathbf{P}^* is the fastest-mixing randomized trajectory, which we proposed for minimizing average age in the information gathering problem. Secondly, the update generation rate for the ground terminal i , depends only on z_{ii} and π_i , which are functions of the first and second moments of the return times to terminal i under trajectory \mathbf{P}^* :

$$\mathbb{E}[H_i] = \frac{1}{\pi_i} \text{ and } \mathbb{E}[H_i^2] = -\frac{1}{\pi_i} + \frac{2z_{ii}}{\pi_i},$$

where H_i denotes the return time to terminal i , starting from i , under the fastest mixing randomized trajectory \mathbf{P}^* . We now bound the performance of this separation principle policy.

Theorem 8: The peak and average age of the separation principle policy is bounded by

$$\frac{A^p}{A_{\mathcal{D}}^p} \leq 4\mathcal{H} + 4\sqrt{\mathcal{H}} + 2 \text{ and } \frac{A^{\text{ave}}}{A_{\mathcal{D}}^{\text{ave}}} \leq 8\mathcal{H} + 8\sqrt{\mathcal{H}} + 4,$$

where \mathcal{H} is the mixing time of the randomized trajectory \mathbf{P}^* .

Proof: We formulate the upper bound age minimization problem and use an approach similar to Lemma 2 and Theorem 6. Proof details are in the technical report [14]. ■

The separation principle policy is factor $O(\mathcal{H})$ peak age and average age optimal. It is worthwhile to note that a similar separation principle policy was established in a completely different setting of scheduling links for age minimization in [7]. Theorem 8 generalizes that result to a graph.

V. SIMULATION RESULTS

We test the performance of our proposed trajectories on three different kinds of mobility graphs: random geometric graphs $\mathcal{G}(n, \frac{2}{\sqrt{n}})$,² grid graphs with diagonal edges, and 3-connected ring or cycle graphs; see Figure 3. We use n to denote the number of ground terminals, namely $n = |V|$. Link weights are picked uniformly at random from the interval $(1, 2]$ in an independent manner. We run our simulations for a total of 50000 time-slots, to get a good estimate of the peak and average age.

We first consider the information gathering problem, and plot peak and average age for all the proposed trajectories of the mobile agent: the Metropolis-Hastings randomized trajectory \mathbf{P}^{mh} , fastest mixing randomized trajectory \mathbf{P}^* , and age-based trajectory. Figure 4 plots peak age as a function of network size n for the random geometric graph $\mathcal{G}(n, 2/\sqrt{n})$. We observe that the peak age for all the three proposed trajectories match. We know from Theorems 3 and 5 that that the two randomized trajectories, namely, the Metropolis-Hastings randomized trajectory \mathbf{P}^{mh} and the fastest mixing randomized trajectory \mathbf{P}^* , are both peak age optimal. Figure 4, therefore, suggests that even the age-based trajectory for the mobile agent is peak age optimal.

In Figure 5 we plot the average age performance of the proposed trajectories, as a function of network size n . Also plotted is the lower bound for average age derived in Theorem 4. We see that the age-based policy is nearly average age optimal, while the fastest mixing randomized trajectory \mathbf{P}^* performs slightly better than the Metropolis-Hastings randomized trajectory \mathbf{P}^{mh} . We compare on a single random instance for each n .

Theorem 6 proved that the fastest mixing randomized trajectory \mathbf{P}^* is at least factor- $8\mathcal{H}$ optimal. Figure 5 validates

²Setting $r = \frac{2}{\sqrt{n}}$ for random geometric graphs ensures connectivity w.h.p.

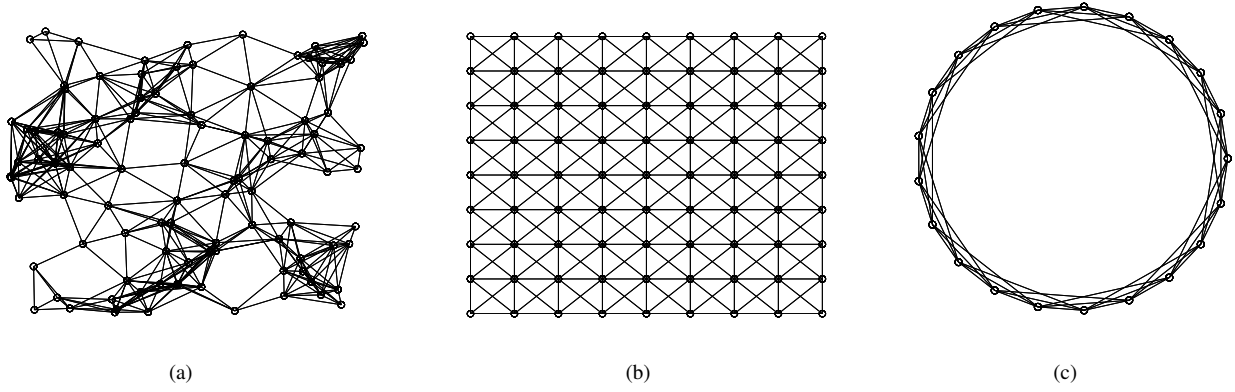


Fig. 3: (a) A random geometric graph with 100 nodes, (b) A grid graph with 81 nodes and diagonal edges, and (c) A 3-connected ring or cycle graph with 21 nodes.

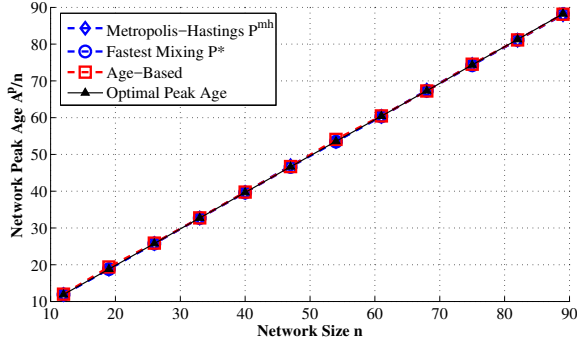


Fig. 4: Information gathering problem in $\mathcal{G}(n, 2/\sqrt{n})$: network peak age as a function of network size n for several proposed trajectories of the mobile agent.

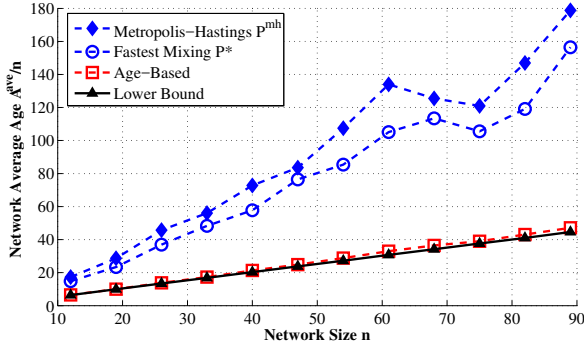


Fig. 5: Information gathering problem in $\mathcal{G}(n, 2/\sqrt{n})$: network average age as a function of network size n for several proposed trajectories of the mobile agent.

this conclusion: for example, for $n = 90$ ground terminals, the average age for the fastest mixing randomized trajectory P^* is approximately a factor 3 away from the lower bound.

In Figure 6 we plot the average age performance for several proposed trajectories, as a function of the network size. The age-based policy, again outperforms the two randomized trajectories, and is nearly optimal. We omit the plot for the ring graph due to space constraints. We observe that the average

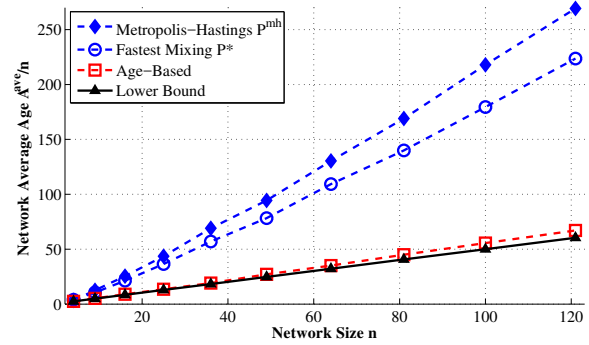


Fig. 6: Information gathering problem in the Grid graph: network average age as a function of network size n for several proposed trajectories of the mobile agent.

age for the fastest mixing randomized trajectory P^* , namely $A^{ave}(P^*)$, is much worse in the ring graph than in the grid graph. This is because the mixing time for the ring graph is much larger than for the grid graph. Similar observation holds in comparing $\mathcal{G}(n, 2/\sqrt{n})$ and the grid graph. Due to space constraints, we omit simulation results for the information dissemination problem.

VI. CONCLUSION

We considered the trajectory planning problem for information gathering and dissemination on graphs using a mobile agent. In the information gathering problem, we showed that a randomized trajectory, namely the fastest-mixing randomized trajectory, is peak age optimal and factor- \mathcal{H} average age optimal. We showed that obtaining an average age optimal trajectory can be NP-hard, while we constructed the peak age optimal trajectory in polynomial time. To improve the average age, we proposed an age-based policy, and showed it to be factor-2 average age optimal, in a symmetric setting. In the information dissemination problem, we proposed a separation principle policy, in which the mobile agent follows the fastest mixing randomized trajectory with a simple rate control. We proved that the separation principle policy is factor- $O(\mathcal{H})$ optimal, in both peak and average age.

REFERENCES

- [1] A. Puri, "A survey of unmanned aerial vehicles (UAV) for traffic surveillance," *Department of computer science and engineering, University of South Florida*, pp. 1–29, 2005.
- [2] T. F. Villa, F. Salimi, K. Morton, L. Morawska, and F. Gonzalez, "Development and validation of a UAV based system for air pollution measurements," *Sensors*, vol. 16, no. 12, p. 2202, 2016.
- [3] D. A. Paley, F. Zhang, and N. E. Leonard, "Cooperative control for ocean sampling: The glider coordinated control system," *IEEE Trans. Control Syst. Technol.*, vol. 16, no. 4, pp. 735–744, 2008.
- [4] A. R. Girard, A. S. Howell, and J. K. Hedrick, "Border patrol and surveillance missions using multiple unmanned air vehicles," in *Proc. CDC*, vol. 1, pp. 620–625, 2004.
- [5] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?," in *Proc. INFOCOM*, pp. 2731–2735, 2012.
- [6] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Trans. Inf. Theory*, vol. 63, pp. 7492–7508, Nov. 2017.
- [7] R. Talak, S. Karaman, and E. Modiano, "Optimizing information freshness in wireless networks under general interference constraints," in *Proc. Mobihoc*, Jun. 2018.
- [8] C. G. Cassandras, X. C. Ding, and X. Lin, "An optimal control approach for the persistent monitoring problem," in *Proc. CDC-ECC*, pp. 2907–2912, 2011.
- [9] S. L. Smith, M. Schwager, and D. Rus, "Persistent monitoring of changing environments using a robot with limited range sensing," in *Proc. ICRA*, pp. 5448–5455, 2011.
- [10] G. D. Celik and E. Modiano, "Dynamic vehicle routing for data gathering in wireless networks," in *Proc. CDC*, pp. 2372–2377, 2010.
- [11] S. Alamdari, E. Fata, and S. L. Smith, "Min-max latency walks: Approximation algorithms for monitoring vertex-weighted graphs," in *Algorithmic Foundations of Robotics X*, pp. 139–155, Springer, 2013.
- [12] J. Liu, X. Wang, B. Bai, and H. Dai, "Age-optimal trajectory planning for UAV-assisted data collection," *arXiv preprint arXiv:1804.09356*, 2018.
- [13] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *Proc. SECON*, pp. 350–358, Jun. 2011.
- [14] Authors, "Age optimal information gathering and dissemination on graphs," *Technical Report*, <https://arxiv.org/abs/1901.02178>, Jan. 2019.
- [15] D. Aldous and J. Fill, "Reversible markov chains and random walks on graphs," 2002.
- [16] J. King, "Conductance and rapidly mixing markov chains," *Technical Report, University of Waterloo*, Mar. 2003.
- [17] N. Ailon, S. Chien, and C. Dwork, "On clusters in markov chains," in *Latin American Symposium on Theoretical Informatics*, pp. 43–55, Springer, 2006.
- [18] S. Boyd, P. Diaconis, and L. Xiao, "Fastest mixing markov chain on a graph," *SIAM Review*, vol. 46, no. 4, pp. 667–689, 2004.
- [19] I. Pak, "Mixing time and long paths in graphs," in *Proc. ACM-SIAM Symposium on Discrete Algorithms*, vol. 6, pp. 321–328, 2002.

APPENDIX

A. Proof of Theorem 3

It suffices to argue that the peak age for any trajectory is lower bounded by $\sum_{i \in V} \frac{w_i}{\pi_i^*}$. Let $H_{k,i}$ to be the k^{th} return time to terminal i . Then the peak age A_i^p is given by

$$A_i^p = \limsup_{T \rightarrow \infty} \frac{\sum_{t=1}^{t=T} A_i(t) \mathbb{1}_{\{m(t)=i\}}}{\sum_{t=1}^{t=T} \mathbb{1}_{\{m(t)=i\}}} = \limsup_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^{k=K} H_{k,i},$$

where K is the total number of returns to ground terminal i over a time horizon T . Now, the fraction of time-slots in

which the mobile agent is at ground terminal i , is given by

$$f_i = \lim_{T \rightarrow \infty} \frac{\sum_{t=1}^{t=T} \mathbb{1}_{\{m(t)=i\}}}{T} = \lim_{K \rightarrow \infty} \frac{K}{\sum_{k=1}^{k=K} H_{k,i}} = \frac{1}{A_i^p},$$

and therefore, $A^p = \sum_{i \in V} w_i A_i^p = \sum_{i \in V} \frac{w_i}{f_i}$. Note that f_i , being the fraction of time-slots the mobile agent is at terminal i , is a distribution over V . Thus, A^p can be lower bounded by

$$A^p = \sum_{i \in V} w_i A_i^p \geq \min_{\{f_i \geq 0, \sum_i f_i = 1\}} \sum_{i \in V} \frac{w_i}{f_i} = \sum_{i \in V} \frac{w_i}{\pi_i^*},$$

where π^* are same as defined in Lemma 1.

B. Proof of Theorem 5

From [18], we know that the fastest mixing, reversible Markov chain on a graph $G(V, E)$ having the stationary distribution π can be found by formulating the following convex program:

$$\begin{aligned} & \underset{\mathbf{P}}{\text{Minimize}} && \|D^{1/2} \mathbf{P} D^{-1/2} - qq^T\|_2, \\ & \text{subject to} && P_{i,j} \geq 0, \forall (i, j) \\ & && \mathbf{P} \mathbf{1} = \mathbf{1}, \\ & && \pi^* \mathbf{P} = \mathbf{P}^T \pi^*, \\ & && P_{i,j} = 0, \forall (i, j) \notin E. \end{aligned} \quad (26)$$

Here $D = \text{diag}(\pi^*)$ and $q = (\sqrt{\pi_1^*}, \sqrt{\pi_2^*}, \dots, \sqrt{\pi_n^*})$. Note that we do not require reversibility, so we can replace the detailed balance constraint $\pi^* \mathbf{P} = \mathbf{P}^T \pi^*$ with the global balance constraint $\pi^* \mathbf{P} = \pi^*$. Also, left and right multiplying $(D^{1/2} \mathbf{P} D^{-1/2} - qq^T)$ by matrices $D^{-1/2}$ and $D^{1/2}$, respectively, does not change the spectral norm; since \mathbf{P} has the same eigen-values as $D^{1/2} \mathbf{P} D^{-1/2}$ and qq^T has the same eigen-values as $D^{-1/2} qq^T D^{1/2}$ [18]. Further, observe that $D^{-1/2} qq^T D^{1/2} = qq^T = \Pi^*$, where $\Pi_{i,j}^* = \pi_i^* \forall i, j \in V$. Thus, the optimization problem reduces to (18).

C. Proof of Theorem 6

Note that the peak age for the fastest-mixing randomized trajectory \mathbf{P}^* is given by $A^p(\mathbf{P}^*) = \sum_{i \in V} \frac{w_i}{\pi_i^*}$, since $\pi^* \mathbf{P}^* = \pi^*$. From Theorem 4, the average age lower bound is given by

$$A_{\text{LB}}^{\text{ave}} = \sum_{i \in V} \frac{1}{2} \left(\frac{w_i}{\pi_i^*} + w_i \right) = \frac{1}{2} A^p(\mathbf{P}^*) + \frac{1}{2} \sum_{i \in V} w_i. \quad (27)$$

To prove the result, it suffices to argue that $A^{\text{ave}}(\mathbf{P}^*)/A_{\text{LB}}^{\text{ave}} \leq 8\mathcal{H}$. From (27) and Lemma 2, we get

$$\frac{A^{\text{ave}}(\mathbf{P}^*)}{A_{\text{LB}}^{\text{ave}}} \leq \frac{4\mathcal{H} A^p(\mathbf{P}^*) + \sum_{i \in V} w_i}{\frac{1}{2} A^p(\mathbf{P}^*) + \frac{1}{2} \sum_{i \in V} w_i} \leq 8\mathcal{H},$$

since \mathcal{H} is always greater than or equal to 1.